

Weather Types and Rainfall over Senegal. Part II: Downscaling of GCM Simulations

VINCENT MORON

CEREGE, Université Aix-Marseille I, Aix en Provence, France, and International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, New York, and Institut Universitaire de France, Paris, France

ANDREW W. ROBERTSON, M. NEIL WARD, AND OUSMANE NDIAYE

International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, New York

(Manuscript received 8 August 2006, in final form 13 April 2007)

ABSTRACT

Four methods of downscaling daily rainfall sequences from general circulation model (GCM) simulations are intercompared over Senegal, using a 13-station network of daily observations during July–September 1961–98. The local scaling method calibrates raw GCM daily rainfall at the closest grid point to a given station so that the climatological distribution of rainfall matches the observed one. The k -nearest neighbor and weather classification schemes resample historical station rainfall observations according to the similarity between the daily wind fields from an ensemble of GCM simulations and a historical library of reanalysis [from 40-yr ECMWF Re-Analysis (ERA-40)] daily wind fields. The nonhomogenous hidden Markov model uses a small set of hidden states to describe the relationship between daily station rainfall observations and low-pass-filtered simulated winds to simulate stochastic sequences of daily rainfall. The four methods are assessed in terms of seasonal statistics of daily rainfall, including seasonal amount, rainfall frequency, and the mean length of wet and dry spells. Verification measures used are mean bias error, interannual anomaly correlation, root-mean-square error, and ranked probability skill score.

The k -nearest neighbor and weather-type classification are shown to perform similarly well in reproducing the mean seasonal cycle, interannual variability of seasonal amount, daily rainfall frequency, as well as the mean length of dry and wet spells, and generally slightly better than the nonhomogeneous hidden Markov model. All three methods are shown to outperform the simple local scaling method. This is due to (i) the ability of the GCM to reproduce remarkably well the mean seasonal cycle and the transitions between weather types defined from reanalysis and (ii) the GCM's moderate-to-strong skill in reproducing the interannual variability of the frequency occurrence of the weather types that strongly influence the interannual variability of rainfall in Senegal. In contrast, the local scaling exaggerates the length of wet and dry spells and reproduces less accurately the interannual variability of the seasonal-averaged amounts, occurrences, and dry/wet spells. This failure is attributed primarily to systematic errors in the GCM's precipitation simulation.

1. Introduction

Decision makers in hydrology, and especially in agronomy, often require daily sequences of weather variables at fine spatial scales to perform local-scale or subregional water or crop modeling studies (i.e., Sultan et al. 2005; Hansen et al. 2006). These time and space scales are smaller than those for which seasonal climate

predictions are issued. Much of the premise of seasonal prediction is based on predictability of large-scale tropical sea surface temperature (SST) anomalies at seasonal lead times, for example, those associated with the El Niño–Southern Oscillation, together with the often-coherent atmospheric response to them. Whether based on empirical relationships or general circulation model (GCM) simulations, seasonal climate predictions are usually expressed in terms of the seasonal amount of rainfall at regional or gridpoint scales (e.g., Goddard et al. 2001). These spatial and temporal scales have the important advantage of filtering out some of the unpredictable noise associated with weather events, local-scale features, measurements errors, etc., but they

Corresponding author address: Vincent Moron, Université d'Aix-Marseille I and CEREGE, UMR 6635 CNRS, Europôle Méditerranéen de l'Arbois, BP 80, 13545 Aix en Provence, France.
E-mail: moron@cerege.fr

do not necessarily fulfill the needs of the wide range of possible end users of these forecasts (Ingram et al. 2002; Hansen et al. 2006).

Downscaling provides a way to utilize GCM seasonal forecasts for local-scale application. This is usually done using either (i) nested regional climate models with a horizontal resolution of 10–50 km (e.g., Sun et al. 2005) or (ii) statistical methods where a transfer function is established between large/seasonal and local/daily scales (e.g., Zorita et al. 1995). Two different statistical approaches can be taken for downscaling to daily rainfall sequences: 1) an “indirect” method using a stochastic weather generator (SWG) to produce the daily sequences of weather conditioned on seasonal/monthly mean variables; 2) a “direct” method where daily sequences are generated from daily GCM outputs. In the latter case, the main principle is (i) to calibrate GCM outputs so that certain properties of the probability density function of observed daily rainfall are reproduced in the simulations (Schmidli et al. 2006; Ines and Hansen 2006), or (ii) involve resampling of historical station records of daily rainfall, based on the similarity between observed and GCM-simulated daily circulation fields (Zorita et al. 1995).

This paper develops and tests two different approaches—weather-type classification and k -nearest neighbor analogs—and compares them with a simple local scaling approach, as well as with a nonhomogeneous hidden Markov model. The weather-type classification scheme is based on the k -means cluster analysis of wind fields over Senegal described in the first part of this study (Moron et al. 2008, hereafter referred to as Part I). Eight weather types were identified using daily wind fields from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) and shown to be related to daily rainfall recorded at 13 Senegalese stations. In this paper we develop a downscaling scheme based on these weather types. The k -nearest neighbor approach (Young 1994; Rajagopalan and Lall 1999; Gangopadhyay et al. 2005) is a generalization of the analog approach (Lorenz 1969; Van den Dool 1989; Toth 1991a,b) in which a large historical archive of maps is searched for the best matches to a target map. Analogs have been used to make predictions some time steps ahead or to establish simultaneous relationships between variables (Zorita and von Storch 1999). The k -nearest neighbor approach extends the sampling of the best match to an ensemble of nearest analogs, allowing ensembles of stochastic daily weather sequences to be generated.

The success of both the weather-type classification and k -nearest neighbor approaches to downscaling is contingent on the ability of the atmospheric GCM to

represent the daily circulation patterns sufficiently well, in terms of their spatial structure and temporal variability. However, it is also evident that days with a similar atmospheric circulation pattern at regional scale may not be similar in rainfall at the local scale. It is expected that generating an ensemble of stochastic realizations helps to better sample the possible outcomes related to a given weather type or atmospheric circulation pattern. This sampling is also crucial for using crop models driven with these daily sequences.

In the local scaling approach, the simulated climatological rainfall occurrence and mean intensity of rainfall on a regular grid are scaled so that they match the observed ones at the closest station (Widmann et al. 2003; Schmidli et al. 2006). The nonhomogeneous hidden Markov model (NHMM) simulates stochastic local-scale daily time series based on a small set of “hidden” states defined from daily rainfall observations at a network of sites (e.g., Hughes and Guttorp 1994; Hughes et al. 1999; Robertson et al. 2004, 2006). It thus provides a classification of the local rainfall patterns that regional-scale atmospheric variability gives rise to. Downscaling is achieved by allowing the Markovian transition probabilities between the states to vary “non-homogeneously” over time according a set of predictors. Each method is used to downscale from an ensemble of GCM simulations made with the ECHAM4.5 model.

The GCM configuration and its mean regional climatology are described in section 2. Section 3 outlines the four downscaling methods. The characteristics of weather types in ECHAM4.5 are described in section 4. The results of the downscaling experiments are described in section 5 using the 13-station daily rainfall observational July–September 1961–98 dataset from the Senegal Weather Service, used in Part I. A summary and discussion are presented in section 6.

2. The GCM simulations

The GCM simulation of circulation fields come from a 24-member ensemble of ECHAM4.5 integrations with observed sea surface temperature (SST) prescribed; each run was made with slightly different initial conditions from January 1950 (Gong et al. 2003; Tippett et al. 2004). The ECHAM4.5 AGCM has been described in detail by Roeckner et al. (1996). The model is a spectral model run at triangular T42 (approximately $2.8^\circ \times 2.8^\circ$) with 19 unevenly spaced levels in the vertical. The daily averages of the zonal and meridional component of the winds at 925, 700, and 200 hPa for the period 1 July 1961 to 30 September 1998 were obtained from the International Research Institute

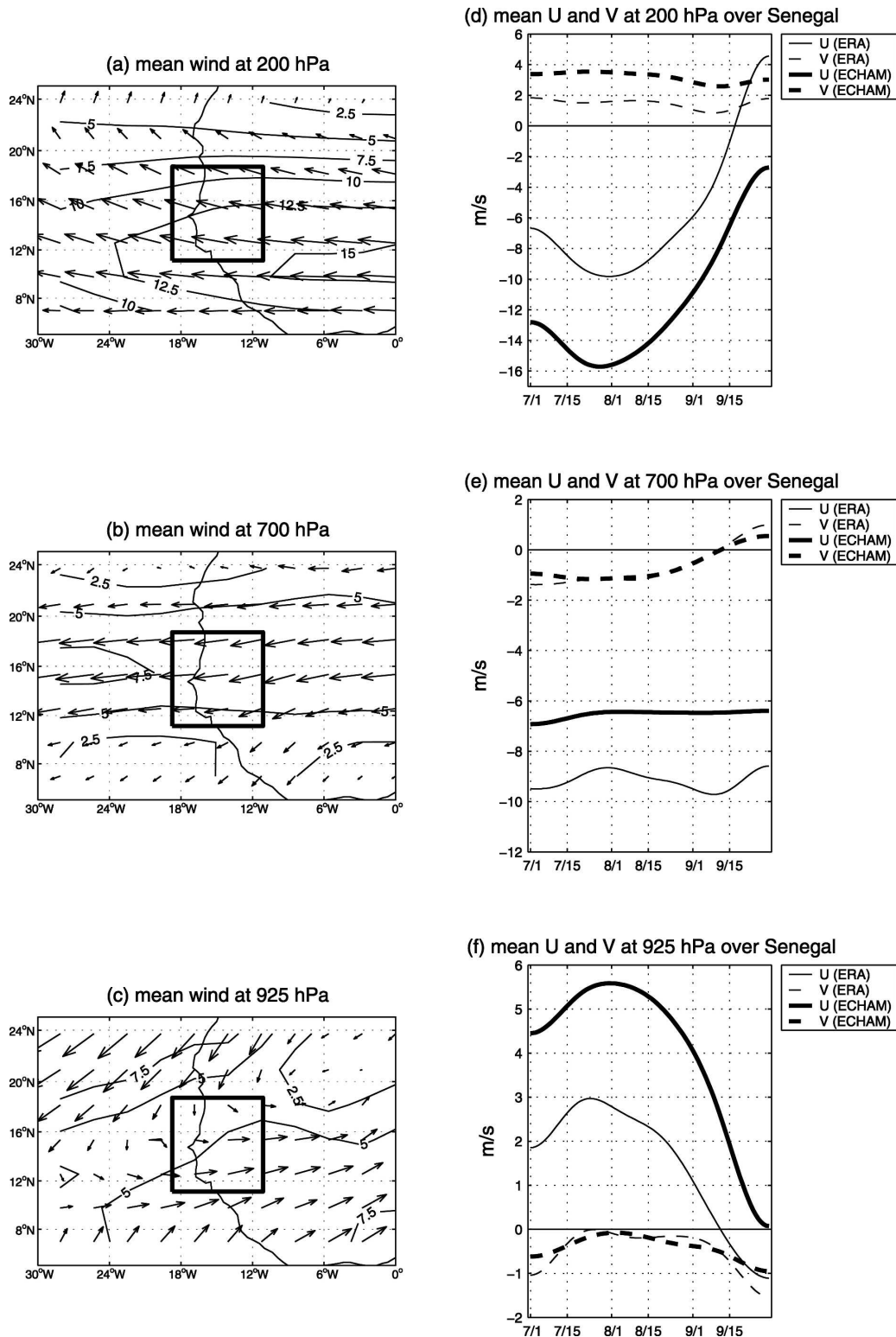


FIG. 1. Mean wind (vector) and speed (contour every 2.5 m s^{-1}) simulated by ECHAM4.5 in July–September 1961–98 at (a) 200, (b) 700, and (c) 925 hPa. Difference between ECHAM4.5 and ERA-40 (Simmons and Gibson 2000) mean winds at (d) 200, (e) 700, and (f) 925 hPa. Seasonal cycle (=mean of daily values low-pass filtered to remove frequencies $>1/30$ cycles day^{-1}) of the zonal component (full line) and meridional (dashed line) of ECHAM4.5 (bold line) and ERA-40 (normal line) averaged over Senegal [i.e., four grid points between 12° and 17°N and 11° and 18°W underlined by a black box in (a)–(c)].

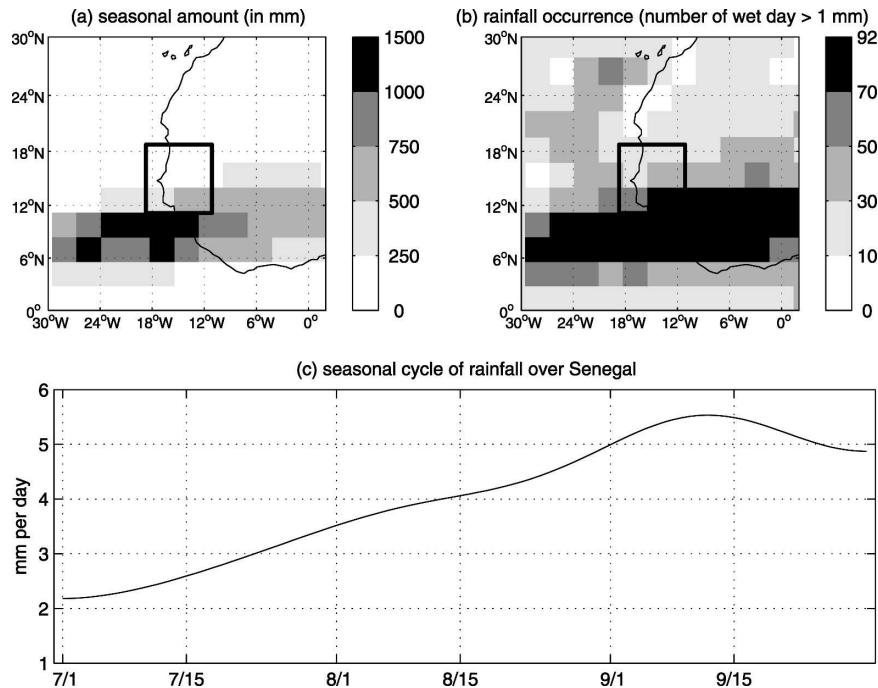


FIG. 2. Mean seasonal (a) amount of rainfall, (b) rainfall occurrence (=number of days receiving >1 mm of rainfall, and (c) seasonal cycle (=mean of daily values low-pass filtered to remove frequencies $>1/30$ cycles day^{-1}) of the rainfall rate (mm day^{-1}) averaged over Senegal [i.e., four grid points between 12° and 17°N and 11° and 18°W underlined by a black box in (a) and (b)] in ECHAM4.5.

for Climate and Society (IRI) data library (<http://ingrid.lidgo.columbia.edu>).

The GCM's ensemble mean climatology of July–September winds at 200, 700, and 925 hPa over West Africa is shown in Fig. 1, together with the mean within-season evolution over Senegal. The GCM reproduces the main features of the West African monsoon including (i) the low-level convergence between the southwesterlies over equatorial Atlantic and tropical West Africa and the northeasterlies to the north (Fig. 1c); (ii) the African easterly jet (AEJ) in the middle troposphere (Fig. 1b); and (iii) the tropical easterly jet (TEJ) in the upper troposphere (Fig. 1a). The phase and amplitude of the seasonal evolution are quite well simulated compared to the ERA-40 reanalysis (Figs. 1d–f). However, the ECHAM zonal winds are too strong at 925 and 200 hPa, roughly by a factor of 2.

Figure 2 shows the GCM's climatological seasonal mean rainfall amount and occurrence frequency, using a 1-mm threshold, together with its mean seasonal evolution of rainfall amount over Senegal. Simulated rainfall is highest over the tropical North Atlantic near 6° – 10°N , 6° – 26°W (Fig. 2a). The seasonal amount is strongly underestimated relative to the rain gauge measurements (see Fig. 1 of Part I), while the number of rainy days is moderately overestimated (i.e., 30–80 wet

days >1 mm in ECHAM versus 15–55 in observations). It is well known that GCMs tend to overestimate the frequency of daily rainfall (“drizzle”) while the mean daily intensity of rainfall is usually underestimated. The simulated seasonal cycle is too flat and the peak is shifted from mid-August to early September in observations (Fig. 1b of Part I) to mid-September in ECHAM (Fig. 1c). This temporal shift is especially clear in central and northern Senegal (not shown).

3. Downscaling methods

a. Local scaling of GCM outputs

The local scaling method (LOC) was proposed by Widmann et al. (2003) for calibrating seasonal GCM rainfall and then extended to the daily time scale by Schmidli et al. (2006) and Ines and Hansen (2006). This simple method provides a benchmark for more sophisticated downscaling methods. Given a station daily rainfall record X , the daily time series of GCM rainfall Y is selected from the GCM grid point closest to each station rainfall record X . Each GCM ensemble member is processed separately. Following Schmidli et al. (2006), the wet-day observed frequency f_{OBS} is taken to be the number of days in X receiving more than 1 mm W_{OBS} of rainfall. The GCM record Y is then sorted

in descending order and the value appearing at rank $f_{\text{OBS}} + 1$ is the threshold W_{GCM} for defining a wet day in the GCM outputs for that station. A scaling factor (S) is computed for calibrating the mean intensity of rainfall on wet days:

$$S = \frac{\bar{X}_{\text{WET}} - W_{\text{OBS}}}{\bar{Y}_{\text{WET}} - W_{\text{GCM}}}, \quad (1)$$

where \bar{X}_{WET} and \bar{Y}_{WET} are the mean observed and simulated wet-day amounts. The calibrated rainfall Y_{scaled} for each station is then given by

$$Y_{\text{scaled}} = S(Y - W_{\text{GCM}}) + W_{\text{OBS}}. \quad (2)$$

The resulting Y_{scaled} below W_{OBS} are set to zero.

To better correct the systematic distortion of the seasonal cycle of rainfall by ECHAM (Fig. 2c), we apply the method to calendar 10-day sequences, rather than months. The resulting threshold values W_{GCM} at each station are plotted in Fig. 3e. The spatial variation of W_{GCM} is consistent with the GCM's mean bias in rainfall occurrence (Fig. 3b), with many values close to 1 mm in the west but larger offsets in the south and southeast where rainfall occurrence is strongly overestimated by ECHAM4.5 (Fig. 3b). Values of W_{GCM} slightly below 1 mm occur when the observed rainfall occurrence is larger than the simulated ones, as for northwest Senegal (Fig. 3e). The scaling factor S is plotted in Fig. 3f and is of the order 2–4 at many inland stations. However, in regions where the GCM's mean intensity is strongly underestimated, particularly in the northwest, S can reach larger values (Fig. 3f).

b. *K*-nearest neighbor and weather-type classification schemes

The *k*-nearest-neighbor analog (KNN) and weather-type classification (WTC) downscaling schemes are both based on assessing the similarity between daily atmospheric circulation patterns in ERA-40 and in the ECHAM4.5 within the region (5° – 25° N, 0° – 30° W) at 925, 700, and 200 hPa. The observed rainfall at the 13 gauge stations on a given day is associated directly with the accompanying ERA-40 circulation pattern on that day.

As described in Part I, the ERA-40 daily wind fields are decomposed into their leading seven empirical orthogonal functions (EOFs), and the set of associated principal components (PC_{era}). The ECHAM4.5 daily wind fields are standardized to zero mean and unit variance at each grid point, but the seasonal cycle is retained. Each GCM ensemble member is then projected linearly onto the ERA-40 EOF space, accounting for 49.2% of the total variance. The resulting GCM time series (PC_{echam}) are scaled so that their mean and variance (for each run) match those of ERA-40; this en-

sures that the GCM wind anomalies can be associated with analogs or weather types identified in the ERA-40 reanalysis (Zorita et al. 1995).

The KNN and WTC schemes are both used to sample daily rainfall vectors conditioned on the similarity between PC_{era} and PC_{echam} . The main difference between the two methods stems from an additional step in weather-type classification, in which the GCM vector PC_{echam} is assigned to the closest ERA-40 weather types (Part I). The next step is to determine the *k*-nearest analogs and closest weather type for each July–September day of each GCM ensemble member. A moving 31-day window centered on each calendar day is used to ensure an accurate simulation of the seasonal cycle of rainfall, so that only calendar dates within 15 days of the target day are used; for the first and last 15 days of the season, the library is reduced accordingly.

Using KNN, there are $n = 31 \times 38 (=1178)$ possible ERA-40 analogs for each GCM day (except at the beginning and the end of the rainy season). The squared Euclidean distances are sorted in ascending order with the smallest one defining the best analog of PC_{echam} . Here we choose $k = \sqrt{n}$ nearest neighbors following Gangopadhyay et al. (2005), yielding $k = 34$. This set of analog days is sampled with replacement using a kernel function (Beersma and Buishand 2003), so that the nearest neighbors are chosen more frequently. The probability that the *j*th closest neighbor is resampled is given by

$$p_j = \frac{1/j}{\sum_{i=1}^k 1/i}, \quad j = 1, \dots, k. \quad (3)$$

For each day of the GCM simulations, we then randomly select 10 ERA-40 days with replacement from the pool of the 34 *k*-nearest neighbor analogs. In the WTC scheme, the ERA-40 days belonging to the weather type assigned to the GCM day are sampled, again within a moving 31-day window, randomly selecting 10 dates with replacement. In both schemes, it is then straightforward to associate each of the chosen ERA-40 dates with the corresponding station rainfall vectors.

c. Overdispersion

The raw application of the above procedures leads to a serious underestimation of the interannual variance of seasonal-averaged station rainfall quantities. This is related to the fact that (i) interannual variability of rainfall over Senegal is only partly explained by changes in weather-type frequency (Part I), and (ii) considering the daily time scale tends to mute the interannual variability, due to the well-known problem of

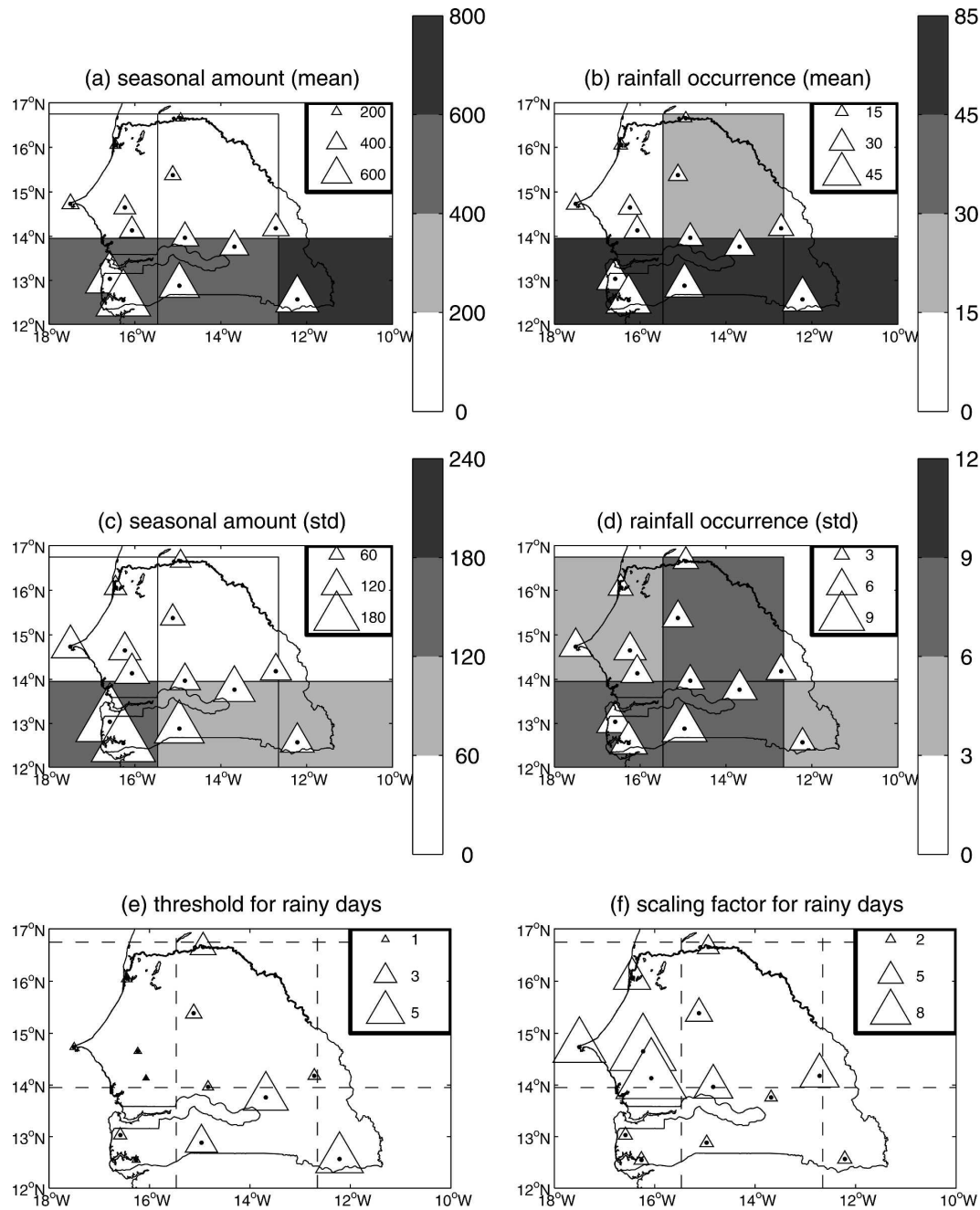


FIG. 3. (a) Long-term mean of the seasonal rainfall (mm) in observations (open triangle) and in ECHAM4.5 (gray shading); (b) long-term mean of the rainfall occurrence (days) in observations (open triangle) and in ECHAM4.5 (gray shading); (c) long-term std dev of the seasonal rainfall (mm) in observations (open triangle) and in ECHAM4.5 (gray shading); (d) long-term std dev of the rainfall occurrence (days) in observations (open triangle) and in ECHAM4.5 (gray shading). All these values are computed with daily rain >1 mm only and the simulated values are the mean of the 24 runs. (e) Mean W_{GCM} (i.e., threshold in mm for defining a wet days in GCM) for each station and (f) dimensionless scaling factor S , to calibrate each daily amounts. In (e) and (f), dashed lines gives the limits of each grid box of ECHAM4.5.

“overdispersion” in stochastic weather generators, which is the underestimation of the interannual variability due to their inability to explicitly take into account this time scale (e.g., Katz and Parlange 1998).

A solution to this problem is to constrain the set of ERA-40 seasons from which the k -nearest neighbors and instances of weather types are drawn. Two criteria are used to constrain this selection based on a principal

component analysis of seasonal-average quantities, namely, (i) the distance between the seasonal values of ECHAM4.5 and ERA-40 winds and (ii) the skill of the GCM's seasonal PCs at simulating their ERA-40 counterparts.

The seasonal averages of ERA-40 zonal and meridional three-level wind fields used in the daily analysis are first reduced to their three leading PCs, which account for 50% of the interannual variance. The GCM's seasonal averages from each ensemble member are then projected linearly onto the ERA-40 EOFs, and resulting time series scaled such that their variances match those of the ERA-40 PCs. The variance of each pair of PC time series is then scaled by the correlation between the seasonal ERA-40 PC and the mean of the 24-member GCM counterparts. Only the two leading PCs, whose skills are 0.76 and 0.29, are retained in the following analyses.

For each simulated season, a kernel function is then used to weight the seasons from which the nearest neighbors and weather-type instances are selected, based on the squared Euclidean distance in the scaled two-dimensional PC subspace. The squared Euclidean distance is first computed between each GCM season and the 38 yr of ERA-40, and the 38 distances then scaled relative to the highest one:

$$d_j = \frac{1/d_j}{1/\max(d_j)}, \quad (4a)$$

where j indicates year. The probability that the j th closest season is resampled is given by

$$p_j = \frac{d_j}{\sum_{j=1}^{38} d_j}. \quad (4b)$$

A sample of 38 seasons of ERA-40 is created with replacement. The seasons that are closer to the target are of course selected multiple times. Then the analyses proceed in exactly the same way as before.

In a few cases (less than 1%), there are no instances of the weather type assigned to a GCM day, within the set of ERA-40 days identified above. When this occurs, the search is relaxed to the whole season rather than the 31-day window, which alleviates the problem.

d. Nonhomogeneous hidden Markov model

As a second benchmark to measure the performance of KNN and WTC methods against, we apply the NHMM. This model is based on the station rainfall records together with a predefined set of predictors that modulate the occurrence of the model's hidden states. The homogeneous HMM factorizes the joint distribution of historical daily rainfall amounts recorded on a

network of stations in terms of a few discrete states, by making two assumptions of conditionality: first, that the rainfall on a given day only depends on the state active on that day, and second, that the state active on a given day depends only on the previous day's state. The latter assumption corresponds to the Markov property, while the fact that the states themselves are not directly observable accounts for the "hidden" in the model description. The NHMM enables downscaling from a set of predictors that then modulate the Markovian transition probabilities between the states "nonhomogeneously" over time. Once the NHMM's parameters have been learned, stochastic simulations of rainfall can be generated at all the stations on the network.

Unlike KNN and WTC, no attempt is made to correct for any overdispersion. We follow the implementation of Robertson et al. (2006), except that the predictors are defined from the GCM zonal and meridional winds to be consistent with KNN and WTC schemes rather than precipitation fields. The winds at 925, 700, and 200 hPa from the GCM simulations are low-pass filtered (cutoff = 10 days), ensemble-averaged, standardized to zero mean and unit variance, and then reduced to their leading two PCs over the domain (5°–25°N, 0°–30°W window), accounting for 66% of the total variance. Both PCs are then standardized to zero mean and unit variance.

We used the cross-validated log-likelihood and the Bayes information criterion (BIC) to evaluate the quality of the fitted HMMs as a function of the number of hidden states, K . The model that best fits the data would have a maximum log-likelihood and a minimum in the BIC. The cross-validated log-likelihood increases substantially from $K = 2$ to $K = 4$ and then levels off, while the BIC reaches a minimum at $K = 5$ (not shown). We choose a value of $K = 5$ in the following, but the downscaling simulation with four states was also tested and found to yield similar results. Cross validation was also used to generate the simulations, with 5 yr withheld at each turn, and 100 stochastic simulations were computed for each season. Cross validation is important in the case of NHMM because, unlike KNN and WTC, a model is built using both the observed rainfall and the GCM simulations. The NHMM involves a logistic regression between the GCM "predictors" variation and the state transitions (Robertson et al. 2004).

4. ECHAM4.5 simulation of the intraseasonal to interannual atmospheric variability

a. Weather types in ECHAM4.5

The ability of ECHAM4.5 to successfully simulate realistic daily rainfall sequences at local scale depends

on various factors, including the ability of ECHAM4.5 to reproduce successfully the atmospheric dynamics at various temporal and spatial scales. There are theoretical limits associated with the spatial coherence of subseasonal rainfall characteristics (Moron et al. 2006, 2007). The framework developed in Part I is used here to examine the ability of the ECHAM4.5 to simulate the weather types identified in ERA-40, the interannual variability of their occurrence, and their subseasonal temporal characteristics.

As described in section 3b, the ECHAM4.5 daily wind anomaly fields at 925, 700, and 200 hPa were projected linearly onto the seven leading EOF patterns of ERA-40, and each simulated day assigned to the closest ERA-40 weather-type centroid identified in Part I. The resulting ECHAM4.5 weather-type composites (not shown) are found to be highly similar to those found in ERA-40 (cf. Figs. 5–8 of Part I), with pattern correlations generally exceeding 0.8. An independent k means classification of the GCM's winds yields similar weather-type patterns. However, this pattern similarity does not imply that the seasonal cycle, transition probabilities, and other temporal characteristics of the GCM's weather types will be close to those found in ERA-40. Figure 4 displays the mean seasonal evolution, transition probabilities, and spell characteristics of the ECHAM4.5 weather types. This figure should be compared with Fig. 4 of Part I. The mean seasonal evolution (Fig. 4a) is reasonably close to that of ERA-40. Nevertheless, there are several differences in the seasonality between ECHAM and ERA-40: (i) WT1 occurrence is underestimated in ECHAM particularly during the early season (Fig. 4a); (ii) the seasonal peak of WT 2 observed in late July in ERA-40 (Fig. 4a of Part I) is slightly underestimated in ECHAM; (iii) WT 7 (WT 8) is overestimated (underestimated) in ECHAM (Fig. 4a) compared to ERA-40. The preferred transitions between WTs are remarkably similar in ECHAM (Fig. 4b) and ERA-40 (Fig. 4b of Part I). In other words, ECHAM reproduces very similar sequences of weather types to ERA-40. The length of homogenous spells (Fig. 4c) is also almost indistinguishable from that of ERA-40 (Fig. 4c of Part I). In summary, the temporal characteristics of the weather types in ECHAM match very closely those found in ERA-40, despite some small differences.

b. Simulated rainfall associated with weather types

It is then interesting to look at the GCM's daily rainfall frequency of occurrence anomalies associated with each weather type (Fig. 5), and to compare these with the station observations in Fig. 9 of Part I. Despite the differences in scale, the negative (positive) rainfall

anomalies are quite reasonably simulated for WTs 1 and 3 (WTs 5 and 6). The negative rainfall anomalies in WT 4, observed mainly over western Senegal (Fig. 9d of Part I) are shifted southeastward by ECHAM (Fig. 5d). Similarly, the positive rainfall anomalies associated with WT 2, which are widespread in observations (Fig. 9b of Part I), are restricted to west and southwest Senegal in ECHAM (Fig. 5b). Weather types 7 and 8 are associated with erroneous rainfall in ECHAM (i.e., strong positive rainfall anomalies over the interior of western Sahel), but this failure is partly related to the GCM's error in the seasonal cycle of rainfall (Fig. 2c). ECHAM shifts the seasonal peak of rainfall toward mid-September, thus artificially inflating the rainfall occurrence of WTs 7 and 8 that peak during that month (Fig. 4a).

c. Potential predictability and skill of the seasonal occurrence of weather types

The last factor examined here is the potential predictability and skill of the GCM's simulation of interannual anomalies in the frequency of occurrence of each weather type. The potential predictability is assessed using the external variance ratio (EVR; Rowell et al. 1995; Zwiers 1996; Rowell 1998), estimated from the 24-member seasonally averaged frequencies of each weather type. The EVR equals 100% if all runs exhibit the same SST-forced interannual variability in terms of the frequency of weather types (Table 1). The skill is assessed in terms of the correlation between the frequency of ERA-40 weather types and the mean of the 24-member ensemble. The most predictable (from SST) and skillful weather type is WT 2, which is interpreted as the long-lasting monsoon surge (Fig. 6a of Part I) and is strongly related to local rainfall (Fig. 11b of Part I). The frequency of occurrences of WTs 3, 1, 7, and 8, which are also related to the interannual variability of local rainfall (Fig. 11b of Part I), are moderately skillful, while WT 5, which is not related to interannual variability of rainfall (Fig. 11b of Part I), is the least reproducible and skillful of the weather types (Table 1).

5. Downscaling results

a. Introduction

The LOC approach yields 24 daily rainfall sequences for each station and season (i.e., one for each run), while the NHMM, KNN, and WTC methods yield 100, 240, and 240 daily sequences, respectively, for each station and season. In the 24 LOC simulations, the spread between simulations is entirely due to differences be-

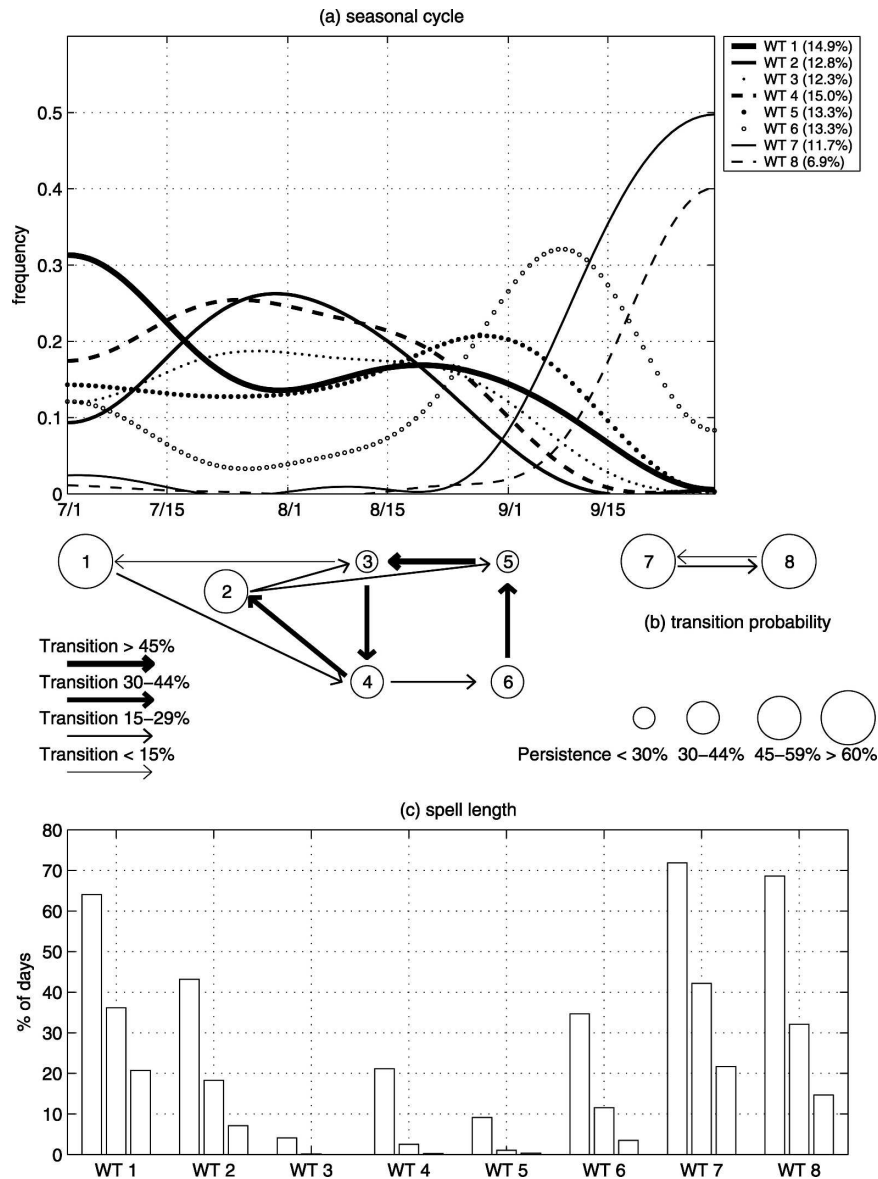


FIG. 4. (a) Seasonal variation of the weather type in ECHAM4.5. The seasonal cycle (=mean of daily value) is smoothed by a low-pass filter retaining periods lower than $1/30$ cycles day^{-1} . (b) Probability transition between the weather types. The arrows indicate the transition that occurs more likely than chance at the one-sided 99% level, and the circles around the weather types are proportional to the one-day persistence. (c) Percentage of days of each type included in homogenous spells lasting at least 3, 5, and 7 days.

tween GCM ensemble members while in KNN and WTC, it also stems from differences between daily rainfall fields within each weather type (WTC) or between two similar atmospheric daily circulation patterns (KNN).

We begin by examining the skill of the four methods for key seasonal rainfall statistics; seasonal amount S , daily rainfall occurrence frequency (daily rainfall > 1 mm) O , and the mean length of spells of consecutive

dry D and wet W days. Skill is assessed in terms of inter-annual anomaly correlations (ACC) between observed and the median of the simulated time series performed for each station and for the standardized anomaly index (SAI) defined as the station average of the standardized anomalies at each station (Figs. 6 and 7). The ranked probability skill score (RPSS) was also computed, using the four stations within or close to the main agricultural region of northwest Senegal (Rigina

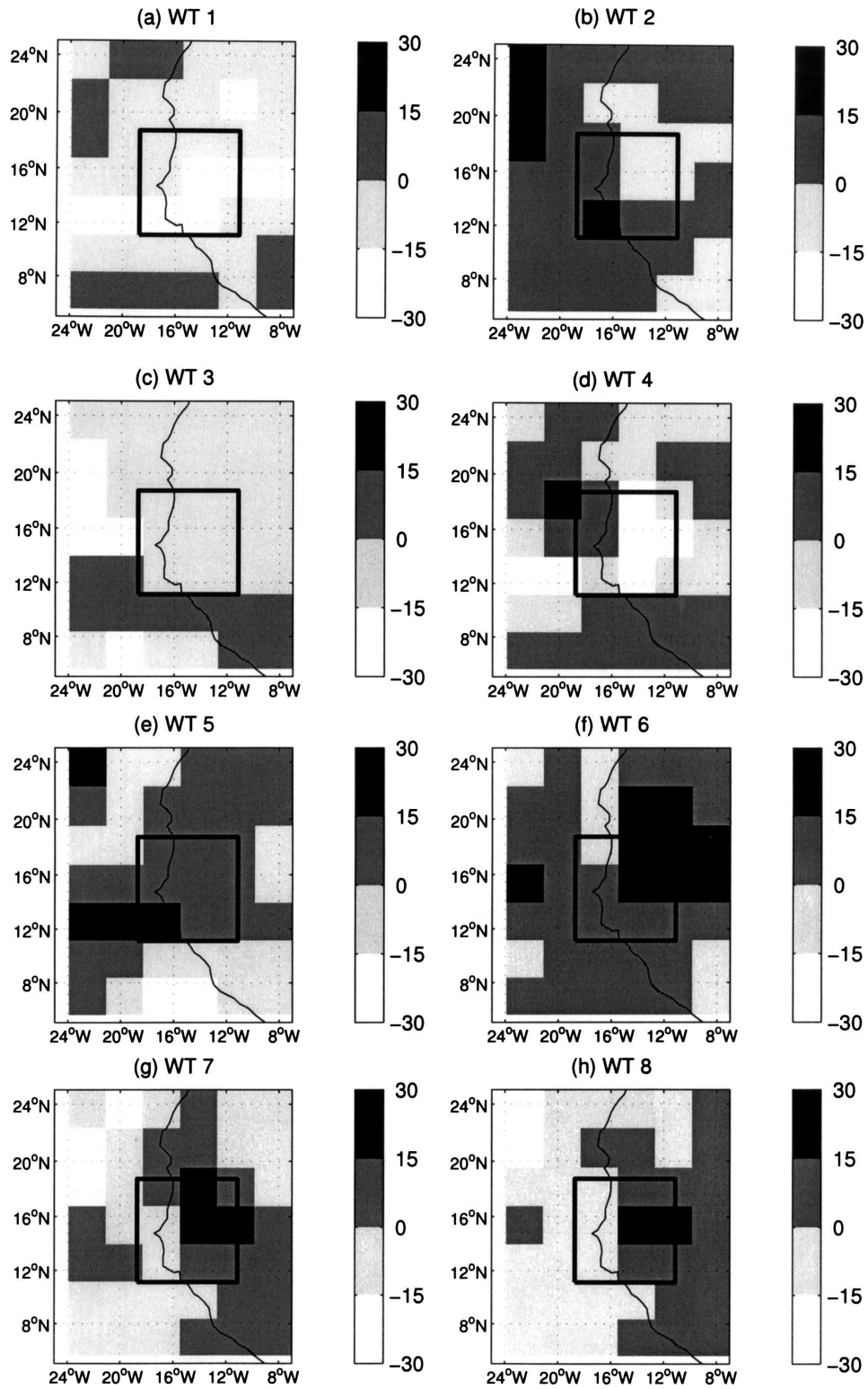


FIG. 5. Rainfall occurrence anomaly (%) simulated for (a) WT 1, (b) WT 2, (c) WT 3, (d) WT 4, (e) WT 5, (f) WT 6, (g) WT 7, and (h) WT 8. The rainfall occurrence is estimated with the days receiving >1 mm of rainfall in ECHAM4.5.

TABLE 1. Potential predictability and skill of ECHAM4.5 weather-type frequency, in terms of the EVR and correlation between ECHAM4.5 ensemble mean and ERA-40.

Weather type	EVR (%)	Skill
1	23.1	0.37
2	38.1	0.64
3	24.6	0.51
4	30.9	0.49
5	7.7	0.03
6	20.4	0.40
7	14.7	0.39
8	17.0	0.44

and Schultz-Rasmussen 2003; Li et al. 2004), that is, Dakar-Yoff, Kaolack, Diourbel, and Kounghel (Fig. 8). The mean bias and root-mean-square error (RMSE) of each method is also computed (Fig. 9, right column).

b. Skill

In terms of anomaly correlation, the KNN and WTC methods yield substantially higher skill than LOC for all four variables, while the NHMM yields only small improvements over LOC. The WTC appears to be slightly better than KNN in terms of ACC. The highest scores are achieved for rainfall occurrence, consistent with the estimates of spatial coherence between rainfall stations and the results obtained from seasonally averaged values by Moron et al. (2006). Significant skill is obtained over most of the country for seasonal amount (Figs. 6e,i) and occurrence (Figs. 6f,j) with highest values for northwest and central Senegal. The differences between stations may be related to physical modulation of the relationships between atmospheric patterns and local rainfall, as well as to sampling variability (Moron et al. 2006, 2007). The skill of the mean length of dry spells is higher for the northern two-thirds of the country (Figs. 6c,g,k,o) while the skill of mean length of wet spell is usually lower with a reversed gradient (Figs. 6d,h,l,p).

The SAI of observed and simulated time series are displayed in Fig. 7 with their ACC given in the right corner of each panel. The SAI illustrates how the KNN, WTC, and NHMM generally improve the simulation of observed variability relative to LOC. They can also sometimes degrade it, such as in 1970–71 where the LOC almost perfectly simulates the SAI seasonal amount (Fig. 7a) and rainfall frequency of occurrence (Fig. 7b) anomalies, while KNN, WTC, and NHMM perform worse for these years. However, LOC performs particularly poorly in 1961, 1976, 1981–83, 1987–90, and 1996 while KNN, WTC, and usually NHMM do a pretty good job for these particular years, at least for seasonal amount and rainfall occurrence (Fig. 7).

The ACCs computed here consider only the median of the simulations and are sensitive to the location of the simulated values with respect to the long-term mean. A probabilistic score such as ranked probability score (Epstein 1969; Murphy 1969) allows the distribution of the simulations to be considered as well as interannual variability of skill. Figure 8 displays the ranked probability skill score (Wilks 1995) for the daily rainfall occurrence frequency of four stations within or near the main agricultural region of Senegal. The RPSS is a categorical probabilistic score and is computed here for quintile categories of the observed time series at each station. The observed and simulated 38-yr time series of frequency of occurrence (=number of wet days >1 mm) are individually standardized to zero mean and unit variance. Four thresholds are computed from ranked observed time series, so that five almost equiprobable classes having respectively 7, 8, 8, 8, and 7 yr are defined. The 24 (LOC), 100 (NHMM), and 240 (WTC and KNN) simulated time series are classified into these classes and RPSS is computed from that. If all members are classified in the observed class, $RPSS = 100\%$. The trivial (or climatological) forecast, with the simulations distributed equally between the five classes gives $RPSS = 0$. The RPSS becomes strongly negative when the simulated values are bounded together far from the observed ones.

At the four individual stations selected in Fig. 8, the median RPSS (i.e., across years) of LOC is almost always weakly negative, except for Diourbel (Fig. 8f). In other words, LOC usually does not outperform on average the climatological forecast. The median RPSS of the three other methods is variable but always positive, with very similar values in the four-station average (Fig. 8j). The frequency of years with $RPSS > 0$ using KNN, WTC, and NHMM exceeds that of LOC (not shown). Despite the improvement of skill provided by KNN, WTC, and NHMM, several years are still simulated wrongly by one or all of the three methods such as 1965, 1970, 1987, 1992, 1995, and 1998, with the number of failures seeming particularly large from 1987.

c. Climatological means and standard deviations

Figure 9 shows a box and whisker plot of the bias in the long-term mean (left column) and standard deviation (middle column), together with the interannual skill in terms of RMSE (right column) of the seasonal amounts (first row), frequency of daily rainfall occurrence (second row), and mean length of dry (third row) and wet spells (fourth row). All these quantities use a threshold of 1 mm to define a wet day. The bias in standard deviation measures the difference between the simulated and observed standard deviations of sea-

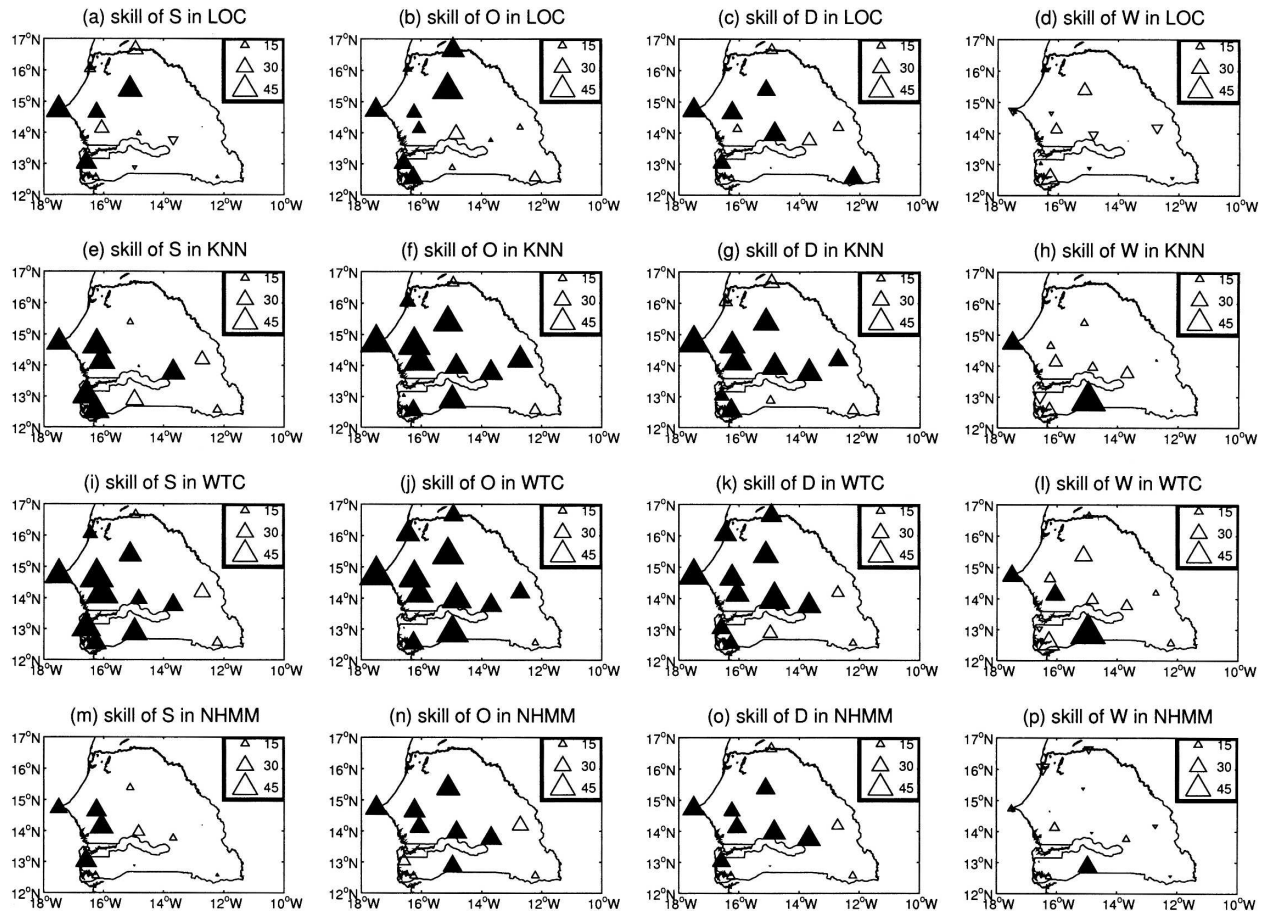


FIG. 6. Correlation $\times 100$ between observed and simulated ($=$ median of the 24, 100, and 240 simulations) for (a)–(d) local scaling, (e)–(h) k -nearest-neighbor analog, (i)–(l) weather-type classification, and (m)–(p) nonhomogeneous hidden Markov model time series for (left) seasonal amount S , (second from the left) frequency of occurrence O , (third from the left) mean length of dry day D , and (right) mean length of wet spell W . Filled triangle shows significant correlations at the two-sided 90% level according to a random-phase test (Janicot et al. 1996; Ebisuzaki 1997).

sonal quantities. The bias and RMSE are computed from each 38-yr simulation at each of the 13 stations and then averaged across the simulations. The upper and lower ends of the box represent the upper and lower quartiles of the 13 station values, and the whiskers show the extent of the rest of the stations excluding any outliers, defined as 1.5 of the interquartile range beyond the upper or lower quartiles; any such outliers are indicated by a cross. Thus the spread in Fig. 9 depicts differences in the bias and RMSE between stations.

By definition, the bias of rainfall occurrence for LOC is zero (Fig. 9d). The bias of seasonal amount is not strictly equal to zero (Fig. 9a) since LOC is performed with a threshold (W_{OBS}) of 1 mm day^{-1} to be consistent with other schemes. The mean bias error for daily rainfall amount and probability of occurrence is also small for KNN, WTC, and NHMM methods, with a bias limited to 8% for the worst station (Fig. 9a). The smallness

of mean bias in KNN and WTC is due to the fact that the observed daily fields are sampled almost equally in the simulations.

The length of dry and wet spells is seriously overestimated in LOC, while KNN, WTC, and NHMM perform almost equally well (Figs. 9g,j). This is related to the fact that the persistence of dry and wet days is strongly overestimated by the GCM and this is reflected directly in LOC. The country average of the persistence of dry and wet days equals, respectively, 0.68 and 0.42 in observations, but 0.73 and 0.51 in LOC. The persistence of dry and wet days is better simulated by KNN (respectively, 0.67 and 0.42), WTC (respectively, 0.67 and 0.39), and NHMM (respectively, 0.67 and 0.43).

LOC also strongly overestimates the amplitude of interannual variability of all four seasonal quantities, while KNN, WTC, and NHMM exhibit small, and

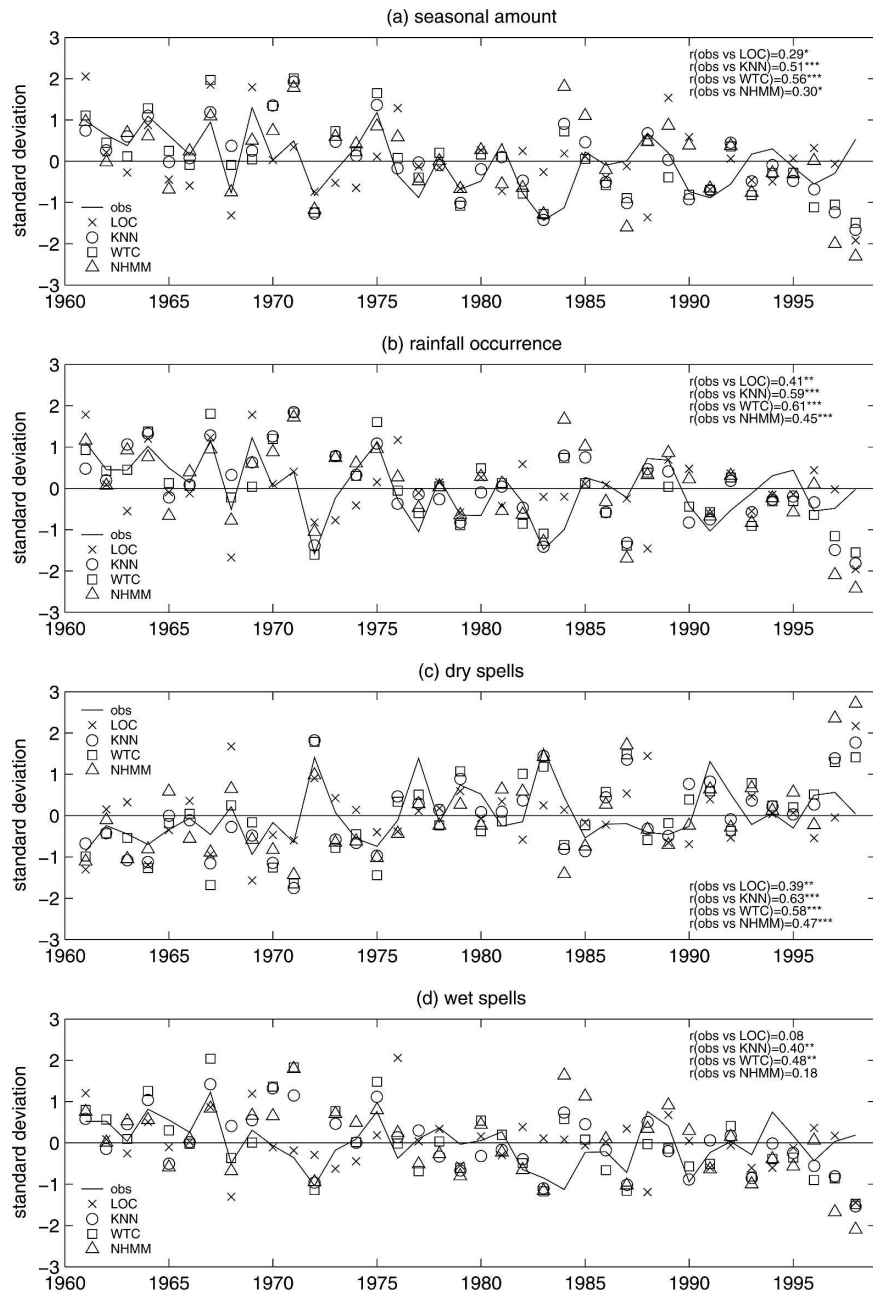


FIG. 7. Observed and simulated standardized anomaly index (=mean of standardized stations) for (a) seasonal amount, (b) occurrence of rainfall, (c) mean length of dry spells, and (d) mean length of wet spells. The ordinates are std devs, and the correlations between observed and simulated SAIs are indicated in the upper or lower right corner of each panel. One, two, and three asterisks indicate significant correlations at the two-sided 90%, 95%, and 99% levels, respectively, according to a random-phase test (Janicot et al. 1996; Ebisuzaki 1997).

rather similar bias in the standard deviations (Figs. 9b,e,h,k). The LOC overestimation for the frequency of occurrence, and the mean length of wet and dry spells, mostly reflects the positive bias of the standard deviations of raw GCM simulations interpolated from the closest grid points (not shown). For seasonal amounts,

this overestimation is also due to the positive scaling parameter S , which tends to increase the mean daily rainfall amounts, particularly over northwest Senegal (Fig. 3f). In contrast to LOC, the two-tier resampling approach of KNN and WTC seems efficient in limiting the overdispersion of the simulations (see section 3c).

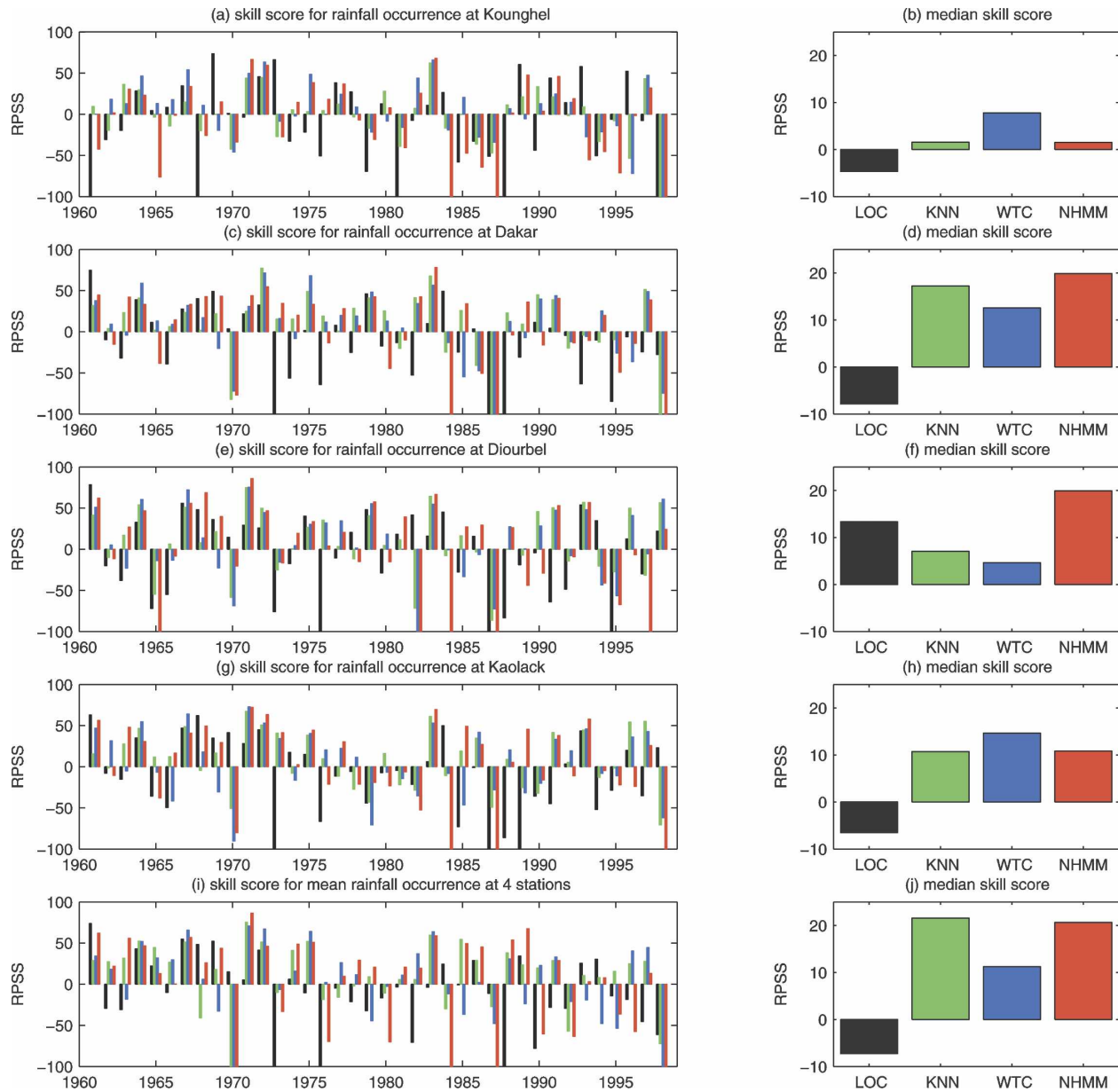


FIG. 8. Individual seasonal ranked probability skill score (in %) for occurrence of rainfall (=frequency of wet day >1 mm) for (a) Kounghele, (c) Dakar-Yoff, (e) Diourbel, (g) Kaolack, and (h) the average of the four stations with the four methods of downscaling (black: local scaling; green: k -nearest-neighbor analog; blue: weather-type classification, and red: nonhomogeneous hidden Markov model). The RPSS $< -100\%$ have been set to -100% . The median RPSS across years for each station and average are displayed [(b), (d), (f), (h), and (j)].

The NHMM also achieves an accurate simulation of the amplitude of interannual variability without any correction for overdispersion.

The RMSE of KNN, WTC, and NHMM are generally smaller than for LOC (Fig. 9). The improvement is especially clear for seasonal amounts (Fig. 9c) and mean length of dry (Fig. 9i) and wet spells (Fig. 9l). The KNN, WTC, and NHMM results are quite similar to each other, with a slight advantage for NHMM (Fig. 9).

Examples of the simulated wet- and dry-spell-length distribution are illustrated in Fig. 10 for the driest (Podor) and wettest (Ziguinchor) stations of the network. In Ziguinchor, the wet spells are longer than the dry spells whereas it is the reverse situation in Podor. The overestimation of the dry and wet spells by LOC is clear for both stations while the KNN, WTC, and NHMM simulations almost match perfectly the observed near-geometric distributions (Fig. 10).

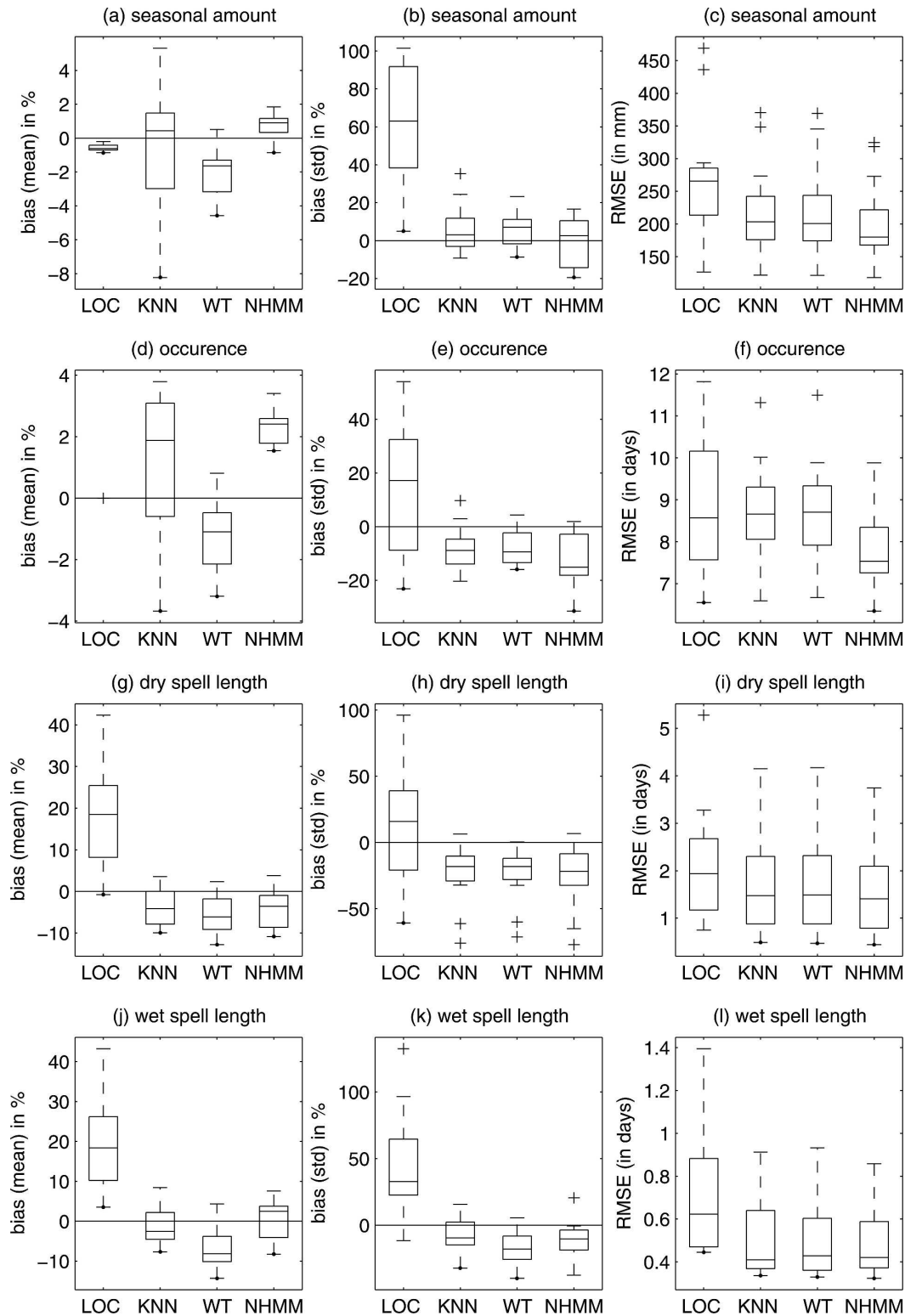


FIG. 9. Box plot of the bias for (left) mean and (middle) std dev, and (right) RMSE of the downscaling daily sequences generated by local scaling, k -nearest-neighbor analog, weather-type classification, and nonhomogeneous hidden Markov model. For each station the bias is estimated from the difference between the whole sample of 24 sequences (LOC), 100 sequences (NHMM), and 240 sequences (KNN and WTC) and the observed time series, then averaged. For each experiment, the lower, middle, and upper bound of the box show the 25th, 50th, and 75th percentiles of the 13-station network, whereas the whiskers show the rest of the data, except for outliers (=crosses) that are above or below 1.5 times the interquartile range from the upper or lower quartiles.

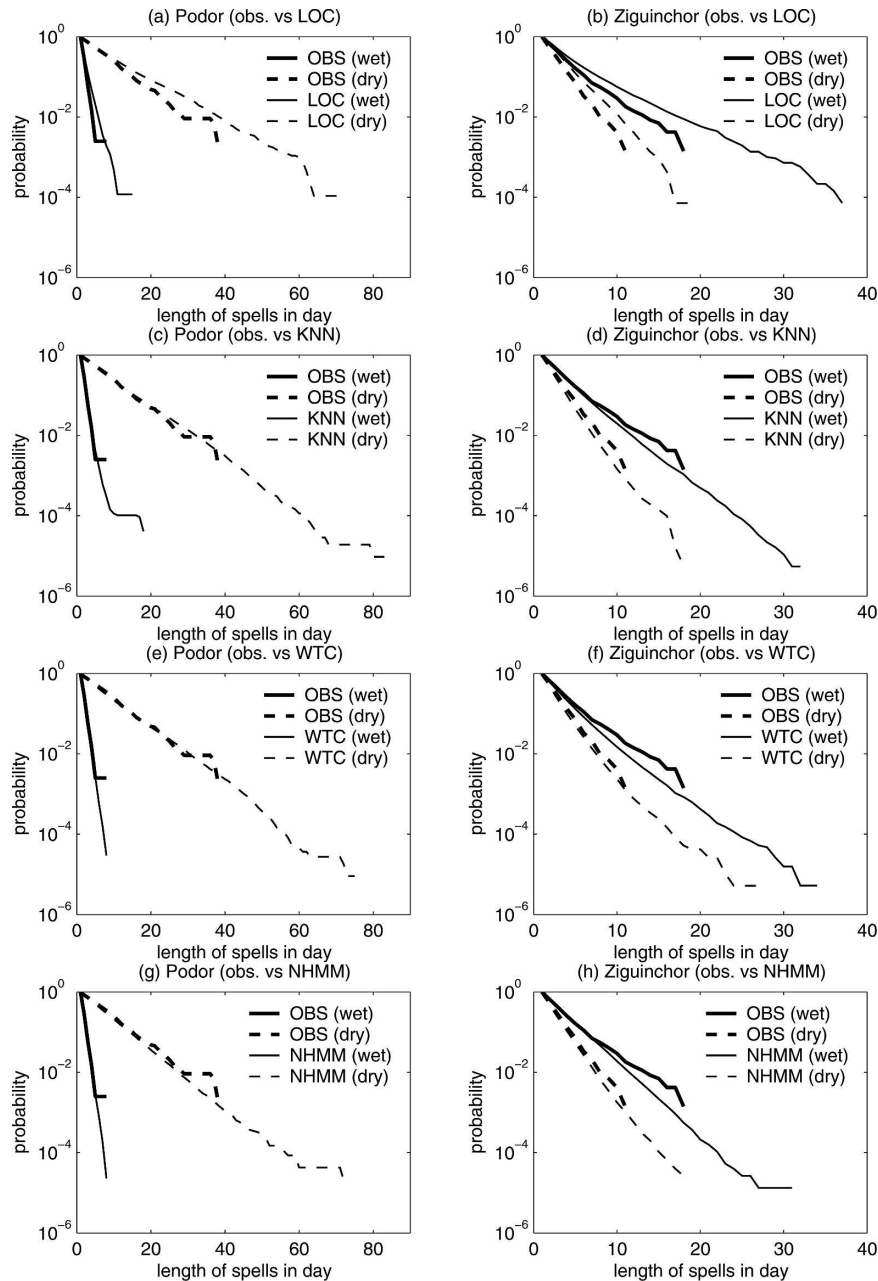


FIG. 10. Distribution of wet (full line) and dry (dashed line) spell length at (left) Podor and (right) Ziguinchor by local scaling, k -nearest-neighbor analog, weather-type classification, and nonhomogeneous hidden Markov model. Thin lines are computed from the downscaled data with the observed data shown as bold line. Plotted is the probability of a spell exceeding a particular duration (days).

6. Summary and discussion

a. Summary

The goal of this paper was to compare four different schemes for downscaling GCM simulations to daily rainfall at the station level. We used a 24-member ensemble of the ECHAM4.5 atmospheric GCM simula-

tions with observed SSTs prescribed, together with a 13-station observed rainfall network over Senegal. The simplest “benchmark” method is a local scaling (LOC) that calibrates the raw daily rainfall from the nearest GCM grid point so that its rainfall occurrence frequency and mean rainfall on wet days match the observed ones over the whole period (Schmidli et al.

2006) (Fig. 3). The LOC is applied to 10-day windows to correct the systematic bias of the long-term climatological seasonal cycle of ECHAM4.5 (Fig. 2c). The k -nearest neighbor (KNN) method searches for the closest circulation analog of each day of the GCM simulation in the ERA-40 library of observed daily circulation fields. The downscaled rainfall is then constructed from the observed precipitation on the days so identified. The weather type classification (WTC) method is based on the weather types presented in Part I. Given a wind field simulated by ECHAM4.5, the method assigns this wind field to an observed weather type determined from ERA-40 and then generates a simulated precipitation value from a randomly chosen observed precipitation field belonging to this weather type. The statistical KNN and WTC methods are both carried out in the subspace of the leading seven EOFs explaining around 52% of the variance of the meridional and zonal components of the daily wind in ERA-40 (Part I). The GCM fields are projected onto these EOFs, ensuring that KNN and WTC are performed on the same base patterns defined from ERA-40. Approximately 50% of the GCM variance is explained by the ERA-40-based patterns. Considering the daily time scale leads to an underestimation of the interannual variance of downscaled rainfall quantities. A two-tier resampling, at seasonal, then at daily time scales, is used to ensure a realistic amplitude of interannual variability. The non-homogeneous hidden Markov model (NHMM) provides another view of the relationships between local-scale rainfall fields and regional-scale atmospheric variability. Five hidden states defined from the 13-station rainfall observations are used to stochastically simulate daily rainfall sequences with predictors taken to be the leading principal components of low-pass-filtered zonal and meridional winds simulated by ECHAM4.5. In the NHMM, the daily predictors influence the transitions between the hidden states.

The KNN, WTC, and NHMM are generally found to perform substantially better than LOC, with the relative skill of the three more complex methods depending on the metric chosen. In terms of anomaly correlation, the KNN and WTC were generally found to perform best in terms of capturing interannual variability of rainfall amount, occurrence, and mean length of dry and wet spells in Senegal (Figs. 6–8). The NHMM performance was generally moderate with LOC being generally the least successful. For example, validating simulated seasonal mean rainfall frequency anomalies across the 13 stations yields a correlation around 0.60 for KNN and WTC, and 0.45 for the NHMM versus 0.41 for LOC (Fig. 7b). However, in terms of RMSE, the NHMM is generally the more skillful (Fig. 9). Con-

sidering RPSS instead of correlation decreases the differences between NHMM, KNN, and WTC (Fig. 8).

In terms of mean bias, the KNN, WTC, and NHMM all performed remarkably well for the climatological mean and interannual standard deviation rainfall anomalies (Fig. 9). LOC has no bias by definition for the seasonal amount and rainfall occurrence but poorly reproduces the mean length of dry and wet spells and usually strongly overestimates the interannual variability. These errors are probably due to systematic biases, particularly the temporal clustering of dry and wet days, in the raw GCM precipitation that are not corrected by the calibration of daily rainfall. In KNN and WTC, the simulation of the seasonal cycle is guaranteed by the fact that both algorithms are performed on a moving 31-day window across the seasonal cycle. In NHMM, a realistic seasonal cycle is reproduced when the predictors have a seasonal cycle as they do here.

b. Discussion

The four downscaling methods used here provide alternative approaches to the stochastic weather generator often used to derive daily sequences of rainfall at local scale from monthly or seasonal regional-scale predictors. LOC is the simplest method, since there is no subjective parameterization, but its application cannot correct for GCM errors in the temporal clustering of wet and dry days and simulated teleconnection patterns. In Senegal, the daily sequences simulated by LOC are too persistent, which could be a serious issue for crop modeling. The LOC simulations also tend to inherit and sometimes exacerbate the GCM's excessively strong interannual variability of simulated precipitation amounts. Nonetheless, LOC provides a measure of the GCM's ability to represent local-scale rainfall variability and defines a benchmark for other methods. The three other methods (i.e., KNN, WTC, and NHMM) are more sophisticated and do indeed lead to better skill, but their success is constrained by three issues: (i) subjective methodological choices, (ii) the strength of the relationship between regional-scale atmospheric variability and the local-scale rainfall, and (iii) the ability of ERA-40 (or any other reanalyses) and the GCM to reproduce the real regional-scale atmospheric variability.

The most important subjective choice in KNN, WTC, and NHMM concerns the predictor selection. The predictors were selected here from the physical point of view, to represent the three distinct vertical levels of the monsoon circulation, that is, the low-level monsoon flow, the AEJ at 700 hPa, and the TEJ at 200 hPa. An alternative approach would be to select the predictors a posteriori from a statistical point of view, to be those

optimizing the rainfall skill at local scale. However, this latter approach would need to be done in a very careful cross-validated fashion, to avoid the serious inherent danger of overfitting. The choice of predictor variables should also be a function of the scales that we expect to be resolved by a GCM, and that are not too sensitive to a particular parameterization. More work is needed to apply our schemes to other GCMs, but the circulation variables chosen here are expected to show less spread between different GCMs than would GCM precipitation fields. The sensitivity of the skill to the choice of number of states (in NHMM) and weather types (in WTC), as well as the proportion of variance in the EOF prefiltering, is very low (not shown).

The second and third issues concern the potential predictability of the regional-scale atmospheric variability at interannual time scale and the physical relationship between the regional-scale daily atmospheric state and local-scale rainfall. At interannual time scale, KNN, WTC, and NHMM benefit from the moderate-to-strong potential predictability and skill of seasonally averaged winds over western Sahel and the nearby northern tropical Atlantic. The success of WTC and KNN relies also on the ability of ECHAM4.5 to reproduce “realistic”—relative to ERA-40—seasonal evolution and sequences of daily atmospheric wind patterns as well as their interannual variability. The similarity found between WTs in ERA-40 (Figs. 5–8 of Part I) and ECHAM4.5 is partly attributable to the projection of the ECHAM4.5 fields onto the ERA-40 EOF patterns, but it is quite remarkable that ECHAM4.5 accurately simulates the transitions (Fig. 4b of Part I and this paper) as well as the mean length of homogenous sequences (Fig. 4c of Part I and this paper) of weather types. In that respect, a possible refinement of KNN and WTC methods could be to consider sequences of days instead of individual days. However, preliminary analyses did not reveal a substantial increase of skill (not shown), probably because the sequences of consecutive wet and dry days are already well captured. At daily time scale, the rainfall–circulation relationship is probably variable across the tropical zone and it is not possible to draw any general conclusion from this study, although the degree of downscaling skill is likely to be constrained by the level of spatial coherence of observed rainfall characteristics (Moron et al. 2006; Part I). The ability of WTC and KNN to generate realistic rainfall sequences is also constrained by the spread amongst rainfall fields associated with close atmospheric patterns (in KNN) and in the same weather types (in WTC). Future works should investigate both sources (i.e., within-GCM ensemble at interannual time

scales and within a pool of similar atmospheric patterns).

The differences between KNN, WTC, and NHMM also deserve some comments. KNN and WTC are very closely related. In some way, KNN uses the information available more efficiently in the sense that it uses the optimal analogs from the whole pool of ERA-40 data while the weather classification approach samples at random from the ERA-40 days that belong to a certain weather type. In contrast to the KNN approach, the WTC approach offers a physical interpretation (Zorita et al. 1995). This physical interpretation is provided by the compression of the unfiltered atmospheric variability into a small subset of weather types that could be related to any local-scale variables, while each analog is intrinsically unique and thus difficult to interpret in the framework of the atmospheric variability. Moreover, this approach also allows an evaluation of the strength and weakness of a current state-of-the-art atmospheric GCM. The NHMM offers a different perspective from WTC and KNN since the NHMM states are largely based on local-scale rainfall, rather than circulation fields in WTC. These rainfall states also enable a physical interpretation, and obviously provide a better summary of the local-scale rainfall—for example, in terms of within-state variance—than do clusters associated with circulation. KNN and WTC are clearly not optimal in that respect, but NHMM is more sensitive to the availability of observed local-scale rainfall. In contrast, KNN and WTC methods are not constrained by the documentation of local-scale rainfall and can be applied for generating daily sequences at a single station and can be easily extended to include other variables such as temperature. It is also conceivable that a single NHMM state could be associated with different atmospheric states, which do not necessarily carry the same amount of potential predictability and skill. On the other hand, the NHMM potentially requires less data to implement because it is not based on daily circulation fields. It also appears to be less sensitive to overdispersion, although this issue requires further studies. Overall, the performance of the three methods are comparable over Senegal, and KNN, WTC, and NHMM can be viewed as complementary methods, offering different perspectives on the complex relationship between regional-scale atmospheric circulation and local-scale rainfall.

Acknowledgments. We thank the editor (David Straus) and the three anonymous reviewers whose constructive comments lead to a substantially improved manuscript. We also thank Benno Blumenthal (IRI, Columbia University) for his help in the extraction of

the 6-h outputs of ECHAM4.5 from the IRI database. ECMWF ERA-40 daily data used in this study have been obtained from the ECMWF data server. The NHMM code was developed by Sergey Kirshner and Padhraic Smyth (University of California, Irvine) and can be obtained online at <http://www.datalab.uci.edu/resources/mvnhmm/>. The Matlab LOC, KNN, and WTC codes can be directly obtained from VM (http://iri.columbia.edu/climate/tools/vincent_matlab_functions/). This paper is funded by a Cooperative Agreement from the National Oceanic and Atmospheric Administration NA05OAR4311004 and by Department of Energy Grant DE-FG02-02ER63413. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its subagencies.

REFERENCES

- Beersma, J. J., and T. A. Buishand, 2003: Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Climate Res.*, **25**, 121–133.
- Ebisuzaki, W., 1997: A method to estimate the statistical significance of a correlation when the data are serially correlated. *J. Climate*, **10**, 2147–2153.
- Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.*, **8**, 985–987.
- Gangopadhyay, S., M. Clark, and B. Rajagopalan, 2005: Statistical downscaling using *K*-nearest neighbors. *Water Resour. Res.*, **41**, W02024, doi:10.1029/2004WR003444.
- Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152.
- Gong, X., A. G. Barnston, and M. N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal precipitation forecasts. *J. Climate*, **16**, 3059–3071.
- Hansen, J. W., A. Challinor, A. V. M. Ines, T. Wheeler, and V. Moron, 2006: Translating climate forecasts into agricultural terms: Advances and challenges. *Climate Res.*, **33**, 27–41.
- Hughes, J. P., and P. Guttorp, 1994: Incorporating spatial dependence and atmospheric data in a model of precipitation. *J. Appl. Meteor.*, **33**, 1503–1515.
- , —, and S. P. Charles, 1999: A non-homogeneous hidden Markov model for precipitation occurrence. *J. Roy. Stat. Soc.*, **48**, 15–30.
- Ines, A. V. M., and J. W. Hansen, 2006: Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteorol.*, **138**, 44–53.
- Ingram, K. T., M. C. Roncoli, and P. H. Kirshen, 2002: Opportunities and constraints for farmers of West Africa to use seasonal precipitation forecasts with Burkina Faso as a case study. *Agric. Syst.*, **74**, 331–349.
- Janicot, S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.*, **23**, 515–518.
- Katz, R. W., and M. B. Parlange, 1998: Overdispersion phenomenon in stochastic modeling of precipitation. *J. Climate*, **11**, 591–601.
- Li, J., J. Lewis, J. Rowland, G. Tappan, and L. L. Tieszen, 2004: Evaluation of land performance in Senegal using multi-temporal NDVI and rainfall series. *J. Arid Environ.*, **59**, 463–480.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646.
- Moron, V., A. W. Robertson, and M. N. Ward, 2006: Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon. Wea. Rev.*, **134**, 3248–3262.
- , —, —, and P. Camberlin, 2007: Spatial coherence of tropical rainfall at the regional scale. *J. Climate*, **20**, 5244–5263.
- , —, —, and O. Ndiaye, 2008: Weather types and rainfall over Senegal. Part I: Observational analysis. *J. Climate*, **21**, 266–287.
- Murphy, A. H., 1969: On the “ranked probability score.” *J. Appl. Meteor.*, **8**, 988–989.
- Rajagopalan, B., and U. Lall, 1999: A *k*-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resour. Res.*, **35**, 3089–3102.
- Rigina, O., and M. Schultz-Rasmussen, 2003: Using trend line and principal component analysis to study vegetation changes in Senegal 1986–1999 from AVHRR NDVI 8 km data. *Dan. J. Geogr.*, **103**, 31–42.
- Robertson, A. W., S. Kirshner, and P. Smyth, 2004: Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *J. Climate*, **17**, 4407–4424.
- , —, —, S. P. Charles, and B. C. Bates, 2006: Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland. *Quart. J. Roy. Meteor. Soc.*, **132**, 519–542.
- Roeckner, E., and Coauthors, 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. Max Planck Institute for Meteorology Rep. 218, Hamburg, Germany, 90 pp.
- Rowell, D. P., 1998: Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J. Climate*, **11**, 109–120.
- , C. K. Folland, K. Maskell, and M. N. Ward, 1995: Variability of summer rainfall over tropical North Africa (1906–92): Observations and modelling. *Quart. J. Roy. Meteor. Soc.*, **121**, 669–704.
- Schmidli, J., C. Frei, and P. L. Vidale, 2006: Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.*, **26**, 679–689.
- Simmons, A. J., and J. K. Gibson, 2000: The ERA-40 project plan. ERA-40 Project Report Series 1, ECMWF, Reading, United Kingdom, 63 pp.
- Sultan, B., C. Baron, M. Dingkuhn, B. Sarr, and S. Janicot, 2005: Agricultural impacts of large-scale variability of the West African monsoon. *Agric. For. Meteorol.*, **128**, 93–110.
- Sun, L., D. F. Moncunill, H. Li, A. D. Moura, and F. de Assis de Souza Filho, 2005: Climate downscaling over Nordeste, Brazil, using the NCEP RSM97. *J. Climate*, **18**, 551–567.
- Tippett, M. K., R. Kleeman, and Y. Tang, 2004: Measuring the potential utility of seasonal climate predictions. *Geophys. Res. Lett.*, **31**, L22201, doi:10.1029/2004GL021575.
- Toth, Z., 1991a: Intercomparison of circulation similarity measures. *Mon. Wea. Rev.*, **119**, 55–64.
- , 1991b: Estimation of atmospheric predictability by circulation analogs. *Mon. Wea. Rev.*, **119**, 65–72.
- Van den Dool, H. M., 1989: A new look at weather forecasting through analogues. *Mon. Wea. Rev.*, **117**, 2230–2247.
- Widmann, M., C. S. Bretherton, and E. P. Salathé Jr., 2003: Sta-

- tistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J. Climate*, **16**, 799–816.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.
- Young, K. C., 1994: A multivariate chain model for simulating climatic parameters from daily data. *J. Appl. Meteor.*, **33**, 661–671.
- Zorita, E., and H. von Storch, 1999: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *J. Climate*, **12**, 2474–2489.
- , J. P. Hughes, D. P. Lettemaier, and H. von Storch, 1995: Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J. Climate*, **8**, 1023–1042.
- Zwiers, F. W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.*, **12**, 825–847.